



why are black women so

why are black women so **angry**

why are black women so **loud**

why are black women so **mean**

why are black women so **attractive**

why are black women so **lazy**

why are black women so **annoying**

why are black women so **confident**

why are black women so **sassy**

why are black women so **insecure**

# ALGORITHMS OF OPPRESSION

HOW SEARCH ENGINES  
REINFORCE RACISM

SAFIYA UMOJA NOBLE

While primarily offering reflection on the effects of search-engine-prioritized content, this research is at the same time intended to bring about a deeper inquiry and a series of strategies that can inform public-policy initiatives focused on connecting Black people to the Internet, in spite of the research that shows that cultural barriers, norms, and power relations alienate Black people from the web.<sup>20</sup> After just over a decade of focus on closing the digital divide,<sup>21</sup> the research questions raised here are meant to provoke a discussion about “what then?” What does it mean to have every Black woman, girl, man, and boy in the United States connected to the web if the majority of them are using a search engine such as Google to access content—whether about themselves or other things—only to find results like those with which I began this introduction? The race to digitize cultural heritage and knowledge is important, but it is often mediated by a search engine for the user who does not know precisely how to find it, much the way a library patron is reliant on deep knowledge and skills of the reference librarian to navigate the vast volumes of information in the library stacks.

### The Importance of Google

Google has become a ubiquitous entity that is synonymous for many everyday users with “the Internet” itself. From serving as a browser of the Internet to handling personal email or establishing Wi-Fi networks and broadband projects in municipalities across the United States, Google, unlike traditional telecommunications companies, has unprecedented access to the collection and provision of data across a variety of platforms in a highly unregulated marketplace and policy environment. We must continue to study the implications of engagement with commercial entities such as Google and what makes them so desirable to consumers, as their use is not without consequences of increased surveillance and privacy invasions and participation in hidden labor practices. Each of these enhances the business model of Google’s parent company, Alphabet, and reinforces its market dominance across a host of vertical and horizontal markets.<sup>22</sup> In 2011, the Federal Trade Commission started looking into Google’s near-monopoly status and market dominance and the harm this could cause consumers. By March 16, 2012, Google was trading on NASDAQ at \$625.04 a share, with a market capitalization of

just over \$203 billion. At the time of the hearings, Google’s latest income statement, for December 2011, showed gross profit at \$24.7 billion. It had \$43.3 billion cash on hand and just \$6.21 billion in debt. Google held 66.2% of the search engine market industry in 2012. Google Search’s profits have only continued to grow, and its holdings have become so significant that the larger company has renamed itself Alphabet, with Google Search as but one of many holdings. By the final writing of this book in August 2017, Alphabet was trading at \$936.38 on NASDAQ, with a market capitalization of \$649.49 billion.

The public is aware of the role of search in everyday life, and people’s opinions on search are alarming. Recent data from tracking surveys and consumer-behavior trends by the comScore Media Metrix consumer panel conducted by the Pew Internet and American Life Project show that search engines are as important to Internet users as email is. Over sixty million Americans engage in search, and for the most part, people report that they are satisfied with the results they find in search engines. The 2005 and 2012 Pew reports on “search engine use” reveal that 73% of all Americans have used a search engine, and 59% report using a search engine every day.<sup>23</sup> In 2012, 83% of search engine users used Google. But Google Search prioritizes its own interests, and this is something far less visible to the public. Most people surveyed could not tell the difference between paid advertising and “genuine” results.

If search is so trusted, then why is a study such as this one needed? The exploration beyond that first simple search is the substance of this book. Throughout the discussion of these and other results, I want to emphasize the main point: there is a missing social context in commercial digital media platforms, and it matters, particularly for marginalized groups that are problematically represented in stereotypical or pornographic ways, for those who are bullied, and for those who are consistently targeted. I use only a handful of illustrative searches to underscore the point and to raise awareness—and hopefully intervention—of how important what we find on the web through commercial search engines is to society.

### Search Results as Power

Search results reflect the values and norms of the search company’s commercial partners and advertisers and often reflect our lowest and

most demeaning beliefs, because these ideas circulate so freely and so often that they are normalized and extremely profitable. Search results are more than simply what is popular. The dominant notion of search results as being both “objective” and “popular” makes it seem as if misogynist or racist search results are a simple mirror of the collective. Not only do problematic search results seem “normal,” but they seem completely unavoidable as well, even though these ideas have been thoroughly debunked by scholars. Unfortunately, users of Google give consent to the algorithms’ results through their continued use of the product, which is largely unavoidable as schools, universities, and libraries integrate Google products into our educational experiences.<sup>24</sup>

Google’s monopoly status,<sup>25</sup> coupled with its algorithmic practices of biasing information toward the interests of the neoliberal capital and social elites in the United States, has resulted in a provision of information that purports to be credible but is actually a reflection of advertising interests. Stated another way, it can be argued that Google functions in the interests of its most influential paid advertisers or through an intersection of popular and commercial interests. Yet Google’s users think of it as a public resource, generally free from commercial interest. Further complicating the ability to contextualize Google’s results is the power of its social hegemony.<sup>26</sup> Google benefits directly and materially from what can be called the “laborertainment”<sup>27</sup> of users, when users consent to freely give away their labor and personal data for the use of Google and its products, resulting in incredible profit for the company.

There are many cases that could be made to show how overreliance on commercial search by the public, including librarians, information professionals, and knowledge managers—all of whom are susceptible to overuse of or even replacement by search engines—is something that we must pay closer attention to right now. Under the current algorithmic constraints or limitations, commercial search does not provide appropriate social, historical, and contextual meaning to already overracialized and hypersexualized people who materially suffer along multiple axes. In the research presented in this study, the reader will find a more meaningful understanding of the kind of harm that such limitations can cause for users reliant on the web as an artifact of both formal and informal culture.<sup>28</sup> In sum, search results play a powerful role in providing fact and authority to those who see them, and as such, they must

be examined carefully. Google has become a central object of study for digital media scholars,<sup>29</sup> due to recognition on these scholars’ parts of the power and impact wielded by the necessity to begin most engagements with social media via a search process and the near universality with which Google has been adopted and embedded into all aspects of the digital media landscape to respond to that need. This work is addressing a gap in scholarship on how search works and what it biases, public trust in search, the relationship of search to information studies, and the ways in which African Americans, among others, are mediated and commodified in Google.

To start revealing some of the processes involved, it is important to think about how results appear. Although one might believe that a query to a search engine will produce the most relevant and therefore useful information, it is actually predicated on a matrix of ways in which pages are hyperlinked and indexed on the web.<sup>30</sup> Rendering web content (pages) findable via search engines is an expressly social, economic, and human project, which several scholars have detailed. These renderings are delivered to users through a set of steps (algorithms) implemented by programming code and then naturalized as “objective.” One of the reasons this is seen as a neutral process is because algorithmic, scientific, and mathematical solutions are evaluated through procedural and mechanistic practices, which in this case includes tracing hyperlinks among pages. This process is defined by Google’s founders, Sergey Brin and Larry Page, as “voting,” which is the term they use to describe how search results move up or down in a ranked list of websites. For the most part, many of these processes have been automated, or they happen through graphical user interfaces (GUIs) that allow people who are not programmers (i.e., not working at the level of code) to engage in sharing links to and from websites.<sup>31</sup>

Research shows that users typically use very few search terms when seeking information in a search engine and rarely use advanced search queries, as most queries are different from traditional offline information-seeking behavior.<sup>32</sup> This front-end behavior of users appears to be simplistic; however, the information retrieval systems are complex, and the formulation of users’ queries involves cognitive and emotional processes that are not necessarily reflected in the system design.<sup>33</sup> In essence, while users use the simplest queries they can in a

search box because of the way interfaces are designed, this does not always reflect how search terms are mapped against more complex thought patterns and concepts that users have about a topic. This disjunction between, on the one hand, users' queries and their real questions and, on the other, information retrieval systems makes understanding the complex linkages between the content of the results that appear in a search and their import as expressions of power and social relations of critical importance.

The public generally trusts information found in search engines. Yet much of the content surfaced in a web search in a commercial search engine is linked to paid advertising, which in part helps drive it to the top of the page rank, and searchers are not typically clear about the distinctions between "real" information and advertising. Given that advertising is a fundamental part of commercial search, using content analysis to make sense of what *actually* is served up in search is appropriate and consistent with the articulation of feminist critiques of the images of women in print advertising.<sup>34</sup> These scholars have shown the problematic ways that women have been represented—as sex objects, incompetent, dependent on men, or underrepresented in the workforce<sup>35</sup>—and the content and representation of women and girls in search engines is consistent with the kinds of problematic and biased ideas that live in other advertising channels. Of course, this makes sense, because Google Search is in fact an advertising platform, not intended to solely serve as a public information resource in the way that, say, a library might. Google creates advertising algorithms, not information algorithms.

To understand search in the context of this book, it is important to look at the description of the development of Google outlined by the former Stanford computer science graduate students and cofounders of the company, Sergey Brin and Larry Page, in "The Anatomy of a Large-Scale Hypertextual Web Search Engine." Their paper, written in graduate school, serves as the architectural framework for Google's PageRank. In addition, it is crucial to also look at the way that citation analysis, the foundational notion behind Brin and Page's idea, works as a bibliometric project that has been extensively developed by library and information science scholars. Both of these dynamics are often misunderstood because they do not account for the complexities of human intervention involved in vetting of information, nor do they pay attention

## Webpage Screenshot

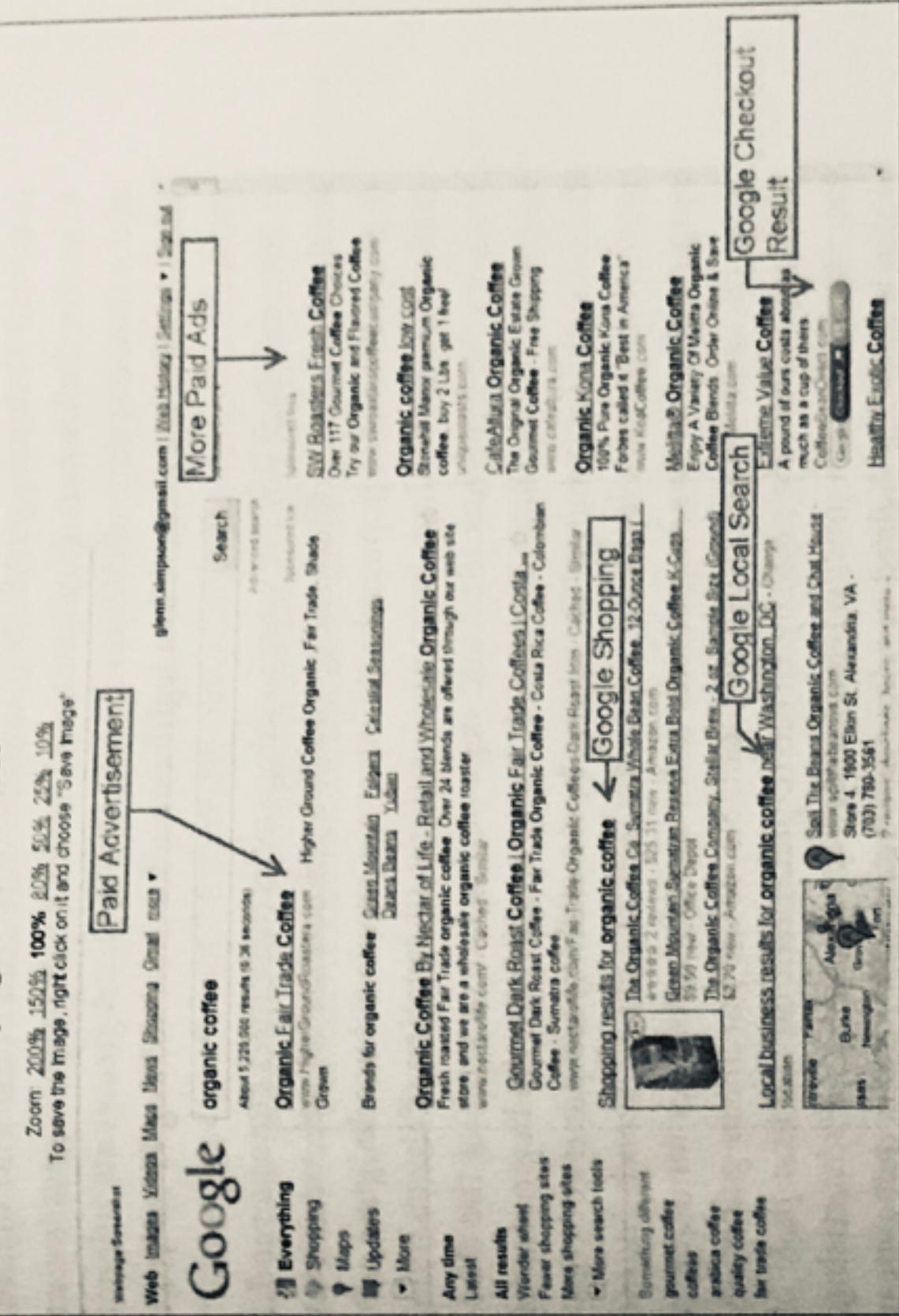


Figure 1.11. Example of Google's prioritization of its own properties in web search.  
Source: Inside Google (2010).

to the relative weight or importance of certain types of information.<sup>36</sup> For example, in the process of citing work in a publication, all citations are given equal weight in the bibliography, although their relative importance to the development of thought may not be equal at all. Additionally, no relative weight is given to whether a reference is validated, rejected, employed, or engaged—complicating the ability to know what a citation actually *means* in a document. Authors who have become so mainstream as not to be cited, such as not attributing modern discussions of class or power dynamics to Karl Marx or the notion of "the individual" to the scholar of the Italian Renaissance Jacob Burckhardt, mean that these intellectual contributions may undergird the framework of an argument but move through works without being cited any longer. Concepts that may be widely understood and accepted ways of knowing are rarely cited in mainstream scholarship, an important dynamic that

Linda Smith, former president of the Association for Information Science and Technology (ASIS&T) and associate dean of the Information School at the University of Illinois at Urbana-Champaign, argues is part of the flawed system of citation analysis that deserves greater attention if bibliometrics are to serve as a legitimating force for valuing knowledge production.

Brin and Page saw the value in using works that others cite as a model for thinking about determining what is legitimate on the web, or at least to indicate what is popular based on many people acknowledging particular types of content. In terms of outright co-optation of the citation, vis-à-vis the hyperlink, Brin and Page were aware of some of the challenges I have described. They were clearly aware from the beginning of the potential for “gaming” the system by advertising companies or commercial interests, a legitimatated process now known as “search engine optimization,” to drive ads or sites to the top of a results list for a query, since clicks on web links can be profitable, as are purchases gained by being vetted as “the best” by virtue of placement on the first page of PageRank. This is a process used for web results, not paid advertising, which is often highlighted in yellow (see figure 1.6). Results that appear not to be advertising are in fact influenced by the advertising algorithm. In contrast to scientific or scholarly citations, which once in print are persistent and static, hyperlinking is a dynamic process that can change from moment to moment.<sup>37</sup> As a result, the stability of results in Google ranking shifts and is prone to being affected by a number of processes that I will cover, primarily search engine optimization and advertising. This means that results shift over time. The results of what is most hyperlinked using Google’s algorithm today will be different at a later date or from the time that Google’s web-indexing crawlers move through the web until the next cycle.<sup>38</sup>

Citation importance is a foundational concept for determining scholarly relevance in certain disciplines, and citation analysis has largely been considered a mechanism for determining whether a given article or scholarly work is important to the scholarly community. I want to revisit this concept because it also has implications for thinking about the legitimation of information, not just citability or popularity. It is also a function of human beings who are engaged in a curation practice, not entirely left to automation. Simply put, if scholars choose to

cite a study or document, they have signaled its relevance; thus, human beings (scholars) are involved in making decisions about a document’s relevance, although all citations in a bibliography do not share the same level of meaningfulness. Building on this concept of credibility through citation, PageRank is what Brin and Page call the greater likelihood that a document is relevant “if there are many pages that point to it” versus “the probability that the random surfer visits a page.”<sup>39</sup> In their research, which led to the development of Google Search, Brin and Page discuss the possibility of monopolizing and manipulating keywords through commercialization of the web search process. Their information-retrieval goal was to deliver the most relevant or very best ten or so documents out of the possible number of documents that could be returned from the web. The resulting development of their search architecture is PageRank—a system that is based on “the objective measure of its citation importance that corresponds well with people’s subjective idea of importance.”<sup>40</sup>

One of the most profound parts of Brin and Page’s work is in appendix A, in which they acknowledge the ways that commercial interests can compromise the quality of search result retrieval. They state, citing Ben Bagdikian, “It is clear that a search engine which was taking money for showing cellular phone ads would have difficulty justifying the page that our system returned to its paying advertisers. For this type of reason and historical experience with other media, we expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers.”<sup>41</sup> Brin and Page outline a clear roadmap for how bias would work in advertising-oriented search and the effects this would have, and they directly suggest that it is in the consumer’s interest not to have search compromised by advertising and commercialism. To some degree, PageRank was intended to be a measure of relevance based on popularity—including what both web surfers and web designers link to from their sites. As with academic citations, Brin and Page decided that citation analysis could be used as a model for determining whether web links could be ranked according to their importance by measuring how much they were back-linked or hyperlinked to or from. Thus, the model for web indexing pages was born. However, in the case of citation analysis, a scholarly author goes through several stages of vetting and credibility testing, such as the peer-review process,

before work can be published and cited. In the case of the web, such credibility checking is not a factor in determining what will be hyperlinked. This was made explicitly clear in the many news reports covering the 2016 U.S. presidential election, where clickbait and manufactured “news” from all over the world clouded accurate reporting of facts on the presidential candidates.

Another example of the shortcomings of removing this human curation or decision making from the first page of results at the top of PageRank, in addition to the results that I found for “black girls,” can be found in the more public dispute over the results that were returned on searches for the word “Jew,” which included a significant number of anti-Semitic pages. As can be seen by Google’s response to the results of a keyword search for “Jew,” Google takes little responsibility toward the ways that it provides information on racial and gendered identities, which are curated in more meaningful ways in scholarly databases. Siva Vaidhyanathan’s 2011 book *The Googlization of Everything (And Why We Should Worry)* chronicles recent attempts by the Jewish community and Anti-Defamation League to challenge Google’s priority ranking to the first page of anti-Semitic, Holocaust-denial websites. So troublesome were these search results that in 2011, Google issued a statement about its search process, encouraging people to use “Jews” and “Jewish people” in their searches, rather than the seemingly pejorative term “Jew”—claiming that the company can do nothing about the word’s co-optation by White supremacist groups (see figure 1.12).

Google, according to its own disclaimer, will only remove pages that are considered unlawful, as is the case in France and Germany, where selling or distributing neo-Nazi materials is prohibited. Without such limits on derogatory, racist, sexist, or homophobic materials, Google allows its algorithm—which is, as we can see, laden with what Diaz calls “sociopolitics”—to stand without debate while protesting its inability to remove pages. As recently as June 27, 2012, Google settled a claim by the French antiracism organization the International League Against Racism over Google’s use of ethnic identity—“Jew”—in association with popular searches.<sup>42</sup> Under French law, racial identity markers cannot be stored in databases, and the auto-complete techniques used in the Google search box link names of people to the word “Jew” on the basis of past user searches. What this recent case points to is another effort to

# Google

## An explanation of our search results

If you recently used Google to search for the word “Jew,” you may have seen results that were very disturbing. We assure you that the views expressed by the sites in your results are not in any way endorsed by Google. We’d like to explain why you’re seeing these results when you conduct this search.

A site’s ranking in Google’s search results relies heavily on computer algorithms using thousands of factors to calculate a page’s relevance to a given query. Sometimes subtleties of language cause anomalies to appear that cannot be predicted. A search for “Jew” brings up one such unexpected result.

If you use Google to search for “Judaism,” “Jewish” or “Jewish people,” the results are informative and relevant. So why is a search for “Jew” different? One reason is that the word ‘Jew’ is often used in an anti-Semitic context. Jewish organizations are more likely to use the word ‘Jewish’ when talking about members of their faith. The word has become somewhat charged linguistically, as noted on websites devoted to Jewish topics such as these:

- <http://www.jewishworldreview.com/cols/jonah081500.asp>

Someone searching for information on Jewish people would be more likely to enter terms like “Judaism,” “Jewish people,” or “Jews” than the single word ‘Jew.’ In fact, prior to this incident, the word ‘Jew’ only appeared about once in every 10 million search queries. Now it’s likely that the great majority of searches on Google for “Jew” are by people who have heard about this issue and want to see the results for themselves.

The beliefs and preferences of those who work at Google, as well as the opinions of the general public, do not determine or impact our search results. Individual citizens and public interest groups do periodically urge us to remove particular links or otherwise adjust search results. Although Google reserves the right to address such requests individually, Google views the comprehensiveness of our search results as an extremely important priority. Accordingly, we do not remove a page from our search results simply because its content is unpopular or because we receive complaints concerning it. We will, however, remove pages from our results if we believe the page (or its site) violates our Webmaster Guidelines, if we believe we are required to do so by law, or at the request of the webmaster who is responsible for the page.

We apologize for the upsetting nature of the experience you had using Google and appreciate your taking the time to inform us about it.

Sincerely,  
The Google Team

P.S. You may be interested in some additional information the Anti-Defamation League has posted about this issue at [http://www.adl.org/rumors/google\\_search\\_tumors.asp](http://www.adl.org/rumors/google_search_tumors.asp). In addition, we call your attention to Google’s search results on this topic.

©2011 Google

Figure 1.12. Explanation of results by Google. Source: [www.google.com/explanation.html](http://www.google.com/explanation.html) (originally available in 2005).

redefine distorted images of people in new media. These cases of distortion, however, continue to accumulate.

The public's as well as the Jewish community's interest in accurate information about Jewish culture and the Holocaust should be enough motivation to provoke a national discussion about consumer harm, to which my research shows we can add other cultural and gender-based identities that are misrepresented in search engines. However, Google's assertion that its search results, though problematic, were computer generated (and thus not the company's fault) was apparently a good-enough answer for the Anti-Defamation League (ADL), which declared, "We are extremely pleased that Google has heard our concerns and those of its users about the offensive nature of some search results and the unusually high ranking of peddlers of bigotry and anti-Semitism."<sup>43</sup> The ADL does acknowledge on its website its gratitude to Sergey Brin, cofounder of Google and son of Russian Jewish immigrants, for his personal letter to the organization and his mea culpa for the "Jew" search-term debacle. The ADL generously stated in its press release about the incident that Google, as a resource to the public, should be forgiven because "until the technical modifications are implemented, Google has placed text on its site that gives users a clear explanation of how search results are obtained. Google searches are automatically determined using computer algorithms that take into account thousands of factors to calculate a page's relevance."<sup>44</sup>

If there is a technical fix, then what are the constraints that Google is facing such that eight years later, the issue has yet to be resolved? A search for the word "Jew" in 2012 produces a beige box at the bottom of the results page from Google linking to its lengthy disclaimer about the results—which remain a mix of both anti-Semitic and informative sites (see figure 1.13). That Google places the responsibility for bad results back on the shoulders of information searchers is a problem, since most of the results that the public gets on broad or open-ended racial and gendered searches are out of their control and entirely within the control of Google Search.

It is important to note that Google has conceded the fact that anti-Semitism as the primary information result about Jewish people is a problem, despite its disclaimer that tries to put the onus for bad results on the searcher. In Germany and France, for example, it is illegal to sell

### Offensive Search Results

[www.google.com/explanation](http://www.google.com/explanation)

We're disturbed about these results as well. Please read our note here.

#### **Searches related to Jew**

- [jew jokes](#)
- [jew watch](#)
- [jew definition](#)
- [jew urban dictionary](#)
- [jewish jokes](#)
- [jew pictures](#)
- [famous jews](#)
- [jew beard](#)

Google >

1 2 3 4 5 6 7 8 9 10

Next

[Advanced search](#)

[Search Help](#)

[Give us feedback](#)

[Google Home](#) [Advertising Programs](#) [Business Solutions](#) [Privacy & Terms](#)

About Google

Figure 1.13. Google's bottom-of-the-page beige box regarding offensive results, which previously took users to "An Explanation of Our Search Results." Source: [www.google.com/explanation](http://www.google.com/explanation) (no longer available).

Nazi memorabilia, and Google has had to put in place filters that ensure online retailers of such are not visible in search results. In 2002, Benjamin Edelman and Jonathan Zittrain at Harvard University's Berkman Center for Internet and Society concluded that Google was filtering its search results in accordance with local law and precluding neo-Nazi organizations and content from being displayed.<sup>45</sup> While this indicates that Google can in fact remove objectionable hits, it is equally troubling, because the company provided search results without informing searchers that information was being deleted. That is to say that the results were presented as factual and complete without mention of omission. Yahoo!, another leading U.S. search engine, was forced into a protracted legal battle in France for allowing pro-Nazi memorabilia to be sold through its search engine, in violation of French law. What these cases point to is that search results are deeply contextual and easily manipulated, rather than objective, consistent, and transparent, and that they can be legitimated only in social, political, and historical context.

The issue of unlawfulness over the harm caused by derogatory results is a question of considerable debate. For example, in the United States, where free speech protections are afforded to all kinds of speech, including hate speech and racist or sexist depictions of people and communities, there is a higher standard of proof required to show harm toward disenfranchised or oppressed people. We need legal protections now more than ever, as automated decision-making systems wield greater power in society.

### Gaming the System: Optimizing and Co-opting Results in Search Engines

Google's advertising tool or optimization product is AdWords. AdWords allows anyone to advertise on Google's search pages and is highly customizable. With this tool, an advertiser can set a maximum amount of money that it wants to spend on a daily basis for advertising. The model for AdWords is that Google will display ads on search pages that it believes are relevant to the kind of search query that is taking place by a user. If a user clicks on an ad, then the advertiser pays. And Google incentivizes advertisers by suggesting that their ads will show up in searches and display, but the advertiser (or Google customer) pays for the ad only when a user (Google consumer) clicks on the advertisement, which is the cost per click (CPC). The advertiser selects a series of "keywords" that it believes closely align with its product or service that it is advertising, and a customer can use a Keyword Estimator tool in order to see how much the keywords they choose to associate with their site might cost. This advertising mechanism is an essential part of how PageRank prioritizes ads on a page, and the association of certain keywords with particular industries, products, and services derives from this process, which works in tandem with PageRank.

In order to make sense of the specific results in keyword searches, it is important to know how Google's PageRank works, what commercial processes are involved in PageRank, how search engine optimization (SEO) companies have been developed to influence the process of moving up results,<sup>46</sup> and how Google bombing<sup>47</sup> occurs on occasion. Google bombing is the practice of excessively hyperlinking to a website (repeatedly coding HTML to link a page to a term or phrase) to cause it to

rise to the top of PageRank, but it is also seen as a type of "hit and run" activity that can deliberately co-opt terms and identities on the web for political, ideological, and satirical purposes. Judit Bar-Ilan, a professor of information science at Bar-Ilan University, has studied this practice to see if the effect of forcing results to the top of PageRank has a lasting effect on the result's persistence, which can happen in well-orchestrated campaigns. In essence, Google bombing is the process of co-opting content or a term and redirecting it to unrelated content. Internet lore attributes the creation of the term "Google bombing" to Adam Mathes, who associated the term "talentless hack" with a friend's website in 2001.

Practices such as Google bombing (also known as Google washing) are impacting both SEO companies and Google alike. While Google is invested in maintaining the quality of search results in PageRank and policing companies that attempt to "game the system," as Brin and Page foreshadowed, SEO companies do not want to lose ground in pushing their clients or their brands up in PageRank.<sup>48</sup> SEO is the process of "using a range of techniques, including augmenting HTML code, web page copy editing, site navigation, linking campaigns and more, in order to improve how well a site or page gets listed in search engines for particular search topics,"<sup>49</sup> in contrast to "paid search," in which the company pays Google for its ads to be displayed when specific terms are searched. A media spectacle of this nature is the case of Senator Rick Santorum, Republican of Pennsylvania, whose website and name were associated with insults in order to drive objectionable content to the top of PageRank.<sup>50</sup> Others who have experienced this kind of co-optation of identity or less-than-desirable association of their name with an insult include former president George W. Bush and the pop singer Justin Bieber.

All of these practices of search engine optimization and Google bombing can take place independently of and in concert with the process of crawling and indexing the web. In fact, being found gives meaning to a website and creates the conditions in which a ranking can happen. Search engine optimization is a major factor in findability on the web. What is important to note is that search engine optimization is a multibillion-dollar industry that impacts the value of specific keywords; that is, marketers are invested in using particular keywords, and keyword combinations, to optimize their rankings.

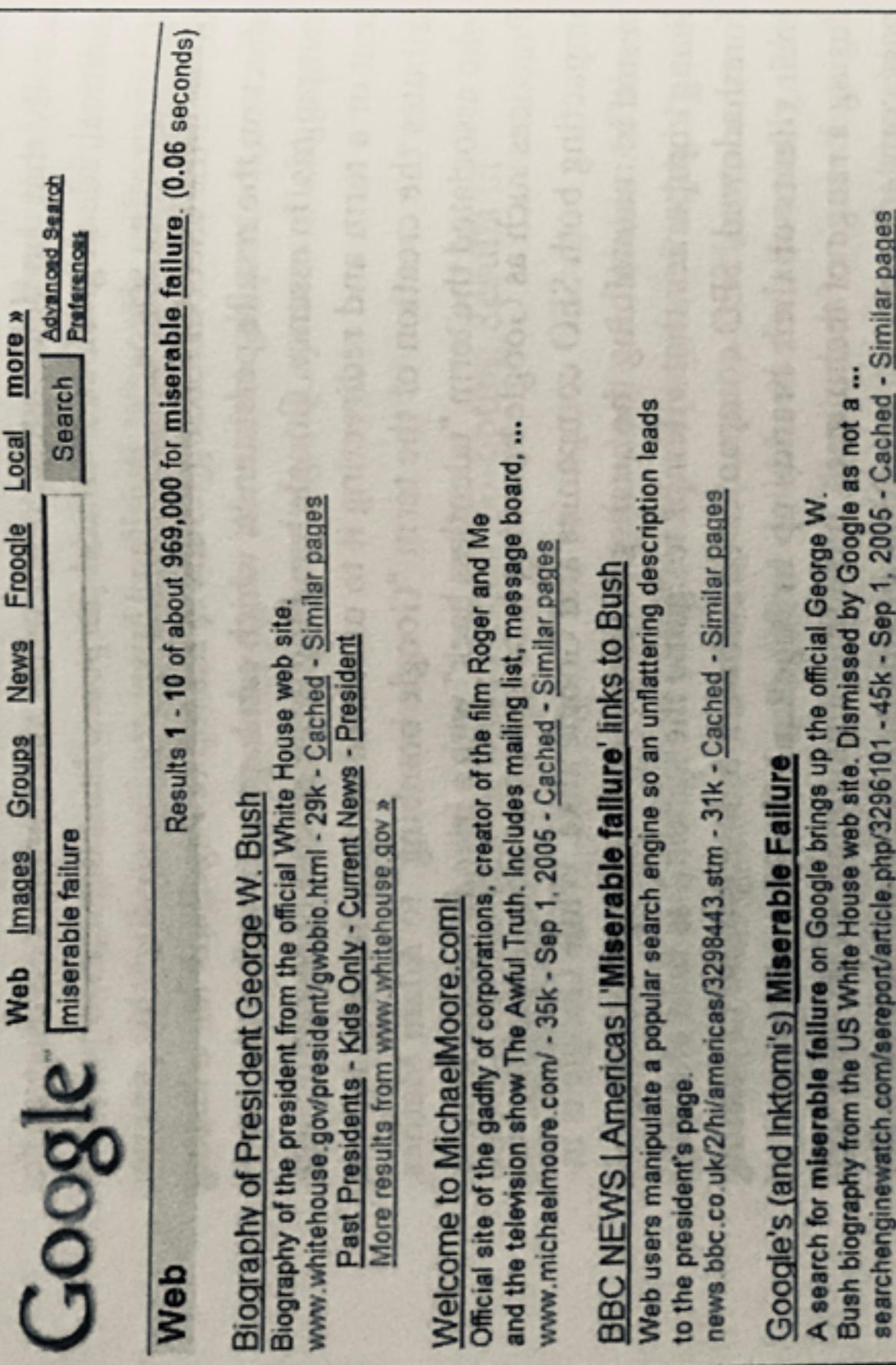


Figure 1.14. Example of a Google bomb on George W. Bush and the search terms “miserable failure,” 2005.

Despite the widespread beliefs in the Internet as a democratic space where people have the power to dynamically participate as equals, the Internet is in fact organized to the benefit of powerful elites,<sup>51</sup> including corporations that can afford to purchase and redirect searches to their own sites. What is most popular on the Internet is not wholly a matter of what users click on and how websites are hyperlinked—there are a variety of processes at play. Max Holloway of *Search Engine Watch* notes, “Similarly, with Google, when you click on a result—or, for that matter, don’t click on a result—that behavior impacts future results. One consequence of this complexity is difficulty in explaining system behavior. We primarily rely on performance metrics to quantify the success or failure of retrieval results, or to tell us which variations of a system work better than others. Such metrics allow the system to be continuously improved upon.”<sup>52</sup> The goal of combining search terms, then, in the context of the landscape of the search engine optimization logic, is only the beginning.

Much research has now been done to dispel the notion that users of the Internet have the ability to “vote” with their clicks and express interest in individual content and information, resulting in democratic practices online.<sup>53</sup> Research shows the ways that political news and information in the blogosphere are mediated and directed such that major news outlets surface to the top of the information pile over less well-known websites and alternative news sites in the blogosphere, to the benefit of elites.<sup>54</sup> In the case of political information seeking, research has shown how Google directs web traffic to mainstream corporate news conglomerates, which increases their ability to shape the political discourse. Google too is a mediating platform that, at least at one moment in time, in September 2011, allowed the porn industry to take precedence in the representations of Black women and girls over other possibilities among at least eleven and a half billion documents that could have been indexed.<sup>55</sup> That moment in 2011 is, however, emblematic of Google’s ongoing dynamic. It has since produced many more problematic results.

As the Federal Communications Commission declares broadband “the new common medium,”<sup>56</sup> the role of search engines is taking on even greater importance to “the widest possible dissemination of information from diverse and antagonistic sources . . . essential to the welfare of the public.”<sup>57</sup> This political economy of search engines and traditional advertisers includes search engine optimization companies that operate in a secondary or gray market (often in opposition to Google). Ultimately, the results we get are about the financial interest that Google or SEOs have in helping their own clients optimize their rankings. In fact, Google is in the business of selling optimization. Extensive critiques of Google have been written on the political economy of search<sup>58</sup> and the way that consolidations in the search engine industry market contribute to the erosion of public resources, in much the way that the media scholars Robert McChesney, former host of nationally syndicated radio show *Media Matters*, and John Nichols, a writer for the *Nation*, critique the consolidation of the mass-media news markets. Others have spoken to the inherent democratizing effect of search engines, such that search is adding to the diversity of political organization and discourse because the public is able to access more information in the marketplace of ideas.<sup>59</sup> Mounting evidence shows that automated decision-making systems are disproportionately harmful to the most vulnerable and the

least powerful, who have little ability to intervene in them—from misrepresentation to prison sentencing to accessing credit and other life impacting formulas.

This landscape of search engines is important to consider in understanding the meaning of search for the public, and it serves as a basis for examining why information quality online is significant. We must trouble the notion of Google as a public resource, particularly as institutions become more reliant on Google when looking for high-quality, contextualized, and credible information. This shift from public institutions such as libraries and schools as brokers of information to the private sector, in projects such as Google Books, for example, is placing previously public assets in the hands of a multinational corporation for private exploitation. Information is a new commodity, and search engines can function as private information enclosures.<sup>60</sup> We need to make more visible the commercial interests that overdetermine what we can find online.

### The Enclosure of the Public Domain through Search Engines

At the same time that search engines have become the dominant portal for information seeking by U.S. Internet users, the rise of commercial mediation of information in those same search engines is further enclosing the public domain. Decreases in funding for public information institutions such as libraries and educational institutions and shifts of responsibility to individuals and the private sector have reframed the ways that the public conceives of what can and should be in the public domain. Yet Google Search is conceived of as a public resource, even though it is a multinational advertising company. These shifts of resources that were once considered public have been impacted by increased intellectual property rights, licensing, and publishing agreements for companies and private individuals in the domain of copyrights, patents, and other legal protections. The move of community-based assets and culture to private hands is arguably a crisis that has rolled back the common good, but there are still possible strategies that can be explored for maintaining what can remain in the public domain. Commercial control over the Internet, often considered a “commons,” has moved it further away from the public through a series of national and international regulations and intellectual and commercial borders that exist in the management of the

network.<sup>61</sup> Beyond the Internet and the control of the network, public information—whether delivered over the web or not—continues to be outsourced to the private sphere, eroding the public information commons that has been a basic tenet of U.S. democracy.

The critical media scholar Herbert Schiller, whose work foreshadowed many of the current challenges in the information and communications landscape, provides a detailed examination of the impact of outsourcing and deregulation in the spheres of communication and public information. His words are still timely: “The practice of selling government (or any) information serves the corporate user well. Ordinarily individual users go to the end of the dissemination queue. Profoundly antidemocratic in its effect, privatizing and/or selling information, which at one time was considered public property, has become a standard practice in recent years.”<sup>62</sup> What this critique shows is that the privatization and commercial nature of information has become so normalized that it not only becomes obscured from view but, as a result, is increasingly difficult to critique within the public domain. The Pew Internet and American Life Project corroborates that the public trusts multinational corporations that provide information over the Internet and that there is a low degree of distrust of the privatization of information.<sup>63</sup> Part of this process of acquiescence to the increased corporatization of public life can be explained by the economic landscape, which is shaped by military-industrial projects such as the Internet that have emerged in the United States,<sup>64</sup> increasing the challenge of scholars who are researching the impact of such shifts in resources and accountability. Molly Niesen at the University of Illinois has written extensively on the loss of public accountability by federal agencies such as the Federal Trade Commission (FTC), which is a major contribution to our understanding of where the public can focus attention on policy interventions.<sup>65</sup> We should leverage her research to think about the FTC as the key agency to manage and intervene in how corporations control the information landscape.

### *The Cultural Power of Algorithms*

The public is minimally aware of these shifts in the cultural power and import of algorithms. In a 2015 study by the Pew Research Center,

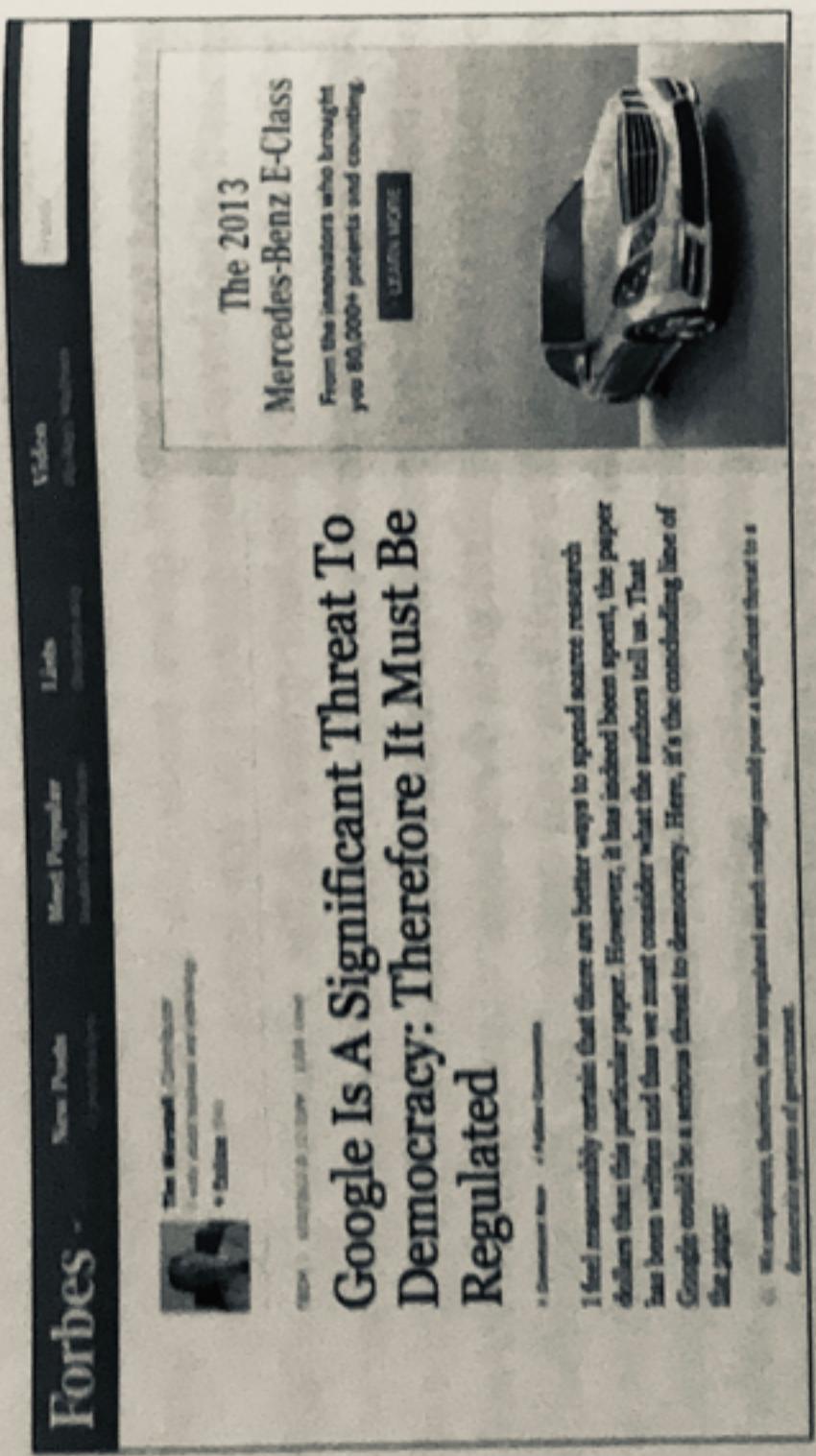


Figure 1.15. *Forbes's* online reporting (and critique) of the Epstein and Robertson study.

"American's Privacy Strategies Post-Snowden," only 34% of respondents who were aware of the surveillance that happens automatically online through media platforms, such as search behavior, email use, and social media, reported that they were shifting their online behavior because of concerns of government surveillance and the potential implications or harm that could come to them.<sup>66</sup> Little of the American public knows that online behavior has more importance than ever. Indeed, Internet-based activities are dramatically affecting our notions of how democracy and freedom work, particularly in the realm of the free flow of information and communication. Our ability to engage with the information landscape subtly and pervasively impacts our understanding of the world and each other.

An example of how information flow and bias in the realm of politics have recently come to the fore can be found in an important new study about how information bias can radically alter election outcomes. The former editor of *Psychology Today* and professor Robert Epstein and Ronald Robertson, the associate director of the American Institute for Behavioral Research and Technology, found in their 2013 study that democracy was at risk because manipulating search rankings could shift voters' preferences, substantially and without their awareness. In their study, they note that the tenor of stories about a candidate in search engine results, whether favorable or unfavorable, dramatically af-

fected the way that people voted. Seventy-five percent of participants were not aware that the search results had been manipulated. The researchers concluded, "The outcomes of real elections—especially tight races—can conceivably be determined by the strategic manipulation of search engine rankings and . . . that the manipulation can be accomplished without people being aware of it. We speculate that unregulated search engines could pose a serious threat to the democratic system of government."<sup>67</sup>

In March 2012, the Pew Internet and American Life Project issued an update to its 2005 "Search Engine Users" study. The 2005 and 2012 surveys tracking consumer-behavior trends from the comScore Media Metrix consumer panel show that search engines are as important to Internet users as email is. In fact, the *Search Engine Use 2012* report suggests that the public is "more satisfied than ever with the quality of search results."<sup>68</sup> Further findings include the following:

- 73% of all Americans have used a search engine, and 59% report using a search engine every day.
  - 83% of search engine users use Google.
- Especially alarming is the way that search engines are increasingly positioned as a trusted public resource returning reliable and credible information. According to Pew, users report generally good outcomes and relatively high confidence in the capabilities of search engines:
- 73% of search engine users say that most or all the information they find as they use search engines is accurate and trustworthy.

Yet, at the same time that search engine users report high degrees of confidence in their skills and trust in the information they retrieve from engines, they have also reported that they are naïve about how search engines work:

- 62% of search engine users are not aware of the difference between paid and unpaid results; that is, only 38% are aware, and only 8% of search engine users say that they can always tell which results are paid or sponsored and which are not.

- In 2005, 70% of search engine users were fine with the concept of paid or sponsored results, but in 2012, users reported that they are not okay with targeted advertising because they do not like having their online behavior tracked and analyzed.
- In 2005, 45% of search engine users said they would stop using search engines if they thought the engines were not being clear about offering some results for pay.
- In 2005, 64% of those who used engines at least daily said search engines are a fair and unbiased source of information; the percentage increased to 66% in 2012.

Users in the 2012 Pew study also expressed concern about personalization:

- 73% reported that they would *not be okay* with a search engine keeping track of searches and using that information to personalize future search results. Participants reported that they feel this to be an invasion of privacy.

and what links we have clicked in order to provide more custom content (a practice that has begun to gather more public attention after Google announced it would use past search practices and link them to users in its privacy policy change in 2012),<sup>70</sup> but search results will also vary depending on whether filters to screen out porn are enabled on computers.<sup>71</sup>

It is certain that information that surfaces to the top of the search pile is not exactly the same for every user in every location, and a variety of commercial advertising, political, social, and economic decisions are linked to the way search results are coded and displayed. At the same time, results are generally quite similar, and complete search personalization—customized to very specific identities, wants, and desires—has yet to be developed. For now, this level of personal-identity personalization has less impact on the variation in results than is generally believed by the public.

### *Losing Control of Our Images and Ourselves in Search*

It is well known that traditional media have been rife with negative or stereotypical images of African American / Black people,<sup>72</sup> and the web as the locus of new media is a place where traditional media interests are replicated. Those who have been inappropriately and unfairly represented in racist and sexist ways in old media have been able to cogently critique those representations and demand expanded representations, protest stereotypes, and call for greater participation in the production of alternative, nonstereotypical or oppressive representations. This is part of the social charge of civil rights organizations such as the Urban League<sup>73</sup> and the National Association for the Advancement of Colored People, which monitor and report on minority misrepresentations, as well as celebrate positive portrayals of African Americans in the media.<sup>74</sup> At a policy level, some civil rights organizations and researchers such as Darnell Hunt, dean of the division of social science and department chair of sociology at UCLA,<sup>75</sup> have been concerned with media representations of African Americans, and mainstream organizations such as Free Press have been active in providing resources about the impact of Facebook track identity and previous searches in order to surface targeted ads for users by analyzing users' web traces. So not only do search engines increasingly remember the digital traces of where we have been

In the context of these concerns, a 2011 study by the researchers Martin Feuz and Matthew Fuller from the Centre for Cultural Studies at the University of London and Felix Stalder from the Zurich University of the Arts found that personalization is not simply a service to users but rather a mechanism for better matching consumers with advertisers and that Google's personalization or aggregation is about actively matching people to groups, that is, categorizing individuals.<sup>69</sup> In many cases, different users are seeing similar content to each other, but users have little ability to see how the platform is attempting to use prior search history and demographic information to shape their results. Personalization is, to some degree, giving people the results they want on the basis of what Google knows about its users, but it is also generating results for viewers to see what Google Search thinks might be good for advertisers by means of compromises to the basic algorithm. This new wave of interactivity, without a doubt, is on the minds of both users and search engine optimizing companies and agencies. Google applications such as Gmail or Google Docs and social media sites such as Facebook track identity and previous searches in order to surface targeted ads for users by analyzing users' web traces. So not only do search engines increasingly remember the digital traces of where we have been

and closing the digital divide.<sup>76</sup> Media advocacy groups that focus on the pornification of women or the stereotyping of people of color might turn their attention toward the Internet as another consolidated media resource, particularly given the evidence showing Google's information and advertising monopoly status on the web.

### Bias in Search

“Traffic Report: How Google Is Squeezing Out Competitors and Muscling Into New Markets,” by ConsumerWatchdog.org’s Inside Google (June 2010), details how Google effectively blocks sites that it competes with and prioritizes its own properties to the top of the search pile (YouTube over other video sites, Google Maps over MapQuest, and Google Images over Photobucket and Flickr). The report highlights the process by which Universal Search is not a neutral and therefore universal process but rather a commercial one that moves sites that buy paid advertising to the top of the pile. Amid these practices, the media, buttressed by an FTC investigation,<sup>77</sup> have suggested that algorithms are not at all unethical or harmful because they are free services and Google has the right to run its business in any way it sees fit. Arguably, this is true, so true that the public should be thoroughly informed about the ways that Google biases information—toward largely stereotypic and decontextualized results, at least when it comes to certain groups of people. Commercial platforms such as Facebook and YouTube go to great lengths to monitor uploaded user content by hiring web content screeners, who at their own peril screen illicit content that can potentially harm the public.<sup>78</sup> The expectation of such filtering suggests that such sites yet content on the Internet on the basis of some objective criteria that indicate that some content is in fact quite harmful to the public. New research conducted by Sarah T. Roberts in the Department of Information Studies at UCLA shows the ways that, in fact, commercial content moderation (CCM, a term she coined) is a very active part of determining what is allowed to surface on Google, Yahoo!, and other commercial text, video, image, and audio engines.<sup>79</sup> Her work on video content moderation elucidates the ways that commercial digital media platforms currently outsource or in-source image and video content filtering to comply with their terms of use

agreements. What is alarming about Roberts’s work is that it reveals the processes by which content is already being screened and assessed according to a continuum of values that largely reflect U.S.-based social norms, and these norms reflect a number of racist and stereotypical ideas that make screening racism and sexism and the abuse of humans in racialized ways “in” and perfectly acceptable, while other ideas such as the abuse of animals (which is also unacceptable) are “out” and screened or blocked from view. She details an interview with one of the commercial content moderators (CCMs) this way:

We have very, very specific itemized internal policies . . . the internal policies are not made public because then it becomes very easy to skirt them to essentially the point of breaking them. So yeah, we had very specific internal policies that we were constantly, we would meet once a week with SecPol to discuss, there was one, blackface is not technically considered hate speech by default. Which always rubbed me the wrong way, so I had probably ten meltdowns about that. When we were having these meetings discussing policy and to be fair to them, they always listened to me, they never shut me up. They didn’t agree, and they never changed the policy but they always let me have my say, which was surprising. (Max Breen, MegaTech CCM Worker).

The MegaTech example is an illustration of the fact that social media companies and platforms make active decisions about what kinds of racist, sexist, and hateful imagery and content they will host and to what extent they will host it. These decisions may revolve around issues of “free speech” and “free expression” for the user base, but on commercial social media sites and platforms, these principles are always counterbalanced by a profit motive; if a platform were to become notorious for being too restrictive in the eyes of the majority of its users, it would run the risk of losing participants to offer to its advertisers. So MegaTech erred on the side of allowing more, rather than less, racist content, in spite of the fact that one of its own CCM team members argued vociferously against it and, by his own description, experienced emotional distress (“meltdowns”) around it.<sup>80</sup>

This research by Roberts, particularly in the wake of leaked reports from Facebook workers who perform content moderation, suggests that people and policies are put in place to navigate and moderate content on the web. Egregious and racist content, content that is highly profitable, proliferates because many tech platforms are interested in attracting the interests and attention of the majority in the United States, not of racialized minorities.

### Challenging Race- and Gender-Neutral Narratives

These explorations of web results on the first page of a Google search also reveal the default identities that are protected on the Internet or are less susceptible to marginalization, pornification, and commodification. The research of Don Heider, the dean of Loyola University Chicago's School of Communication, and Dustin Harp, an assistant professor in the Department of Communication at the University of Texas, Arlington, shows that even though women constitute just slightly over half of Internet users, women's voices and perspectives are not as loud and do not have as much impact online as those of men. Their work demonstrates how some users of the Internet have more agency and can dominate the web, despite the utopian and optimistic view of the web as a socially equalizing and democratic force.<sup>81</sup> Recent research on the male gaze and pornography on the web argue that the Internet is a communications environment that privileges the male, pornographic gaze and marginalizes women as objects.<sup>82</sup> As with other forms of pornographic representations, pornography both structures and reinforces the domination of women, and the images of women in advertising and art are often "constructed for viewing by a male subject,"<sup>83</sup> reminiscent of the journalist and producer John Berger's canonical work *Ways of Seeing*, which describes this objectification in this way: "Women are depicted in a quite different way from men—not because the feminine is different from the masculine—but because the 'ideal' spectator is always assumed to be male and the image of the woman is designed to flatter him."<sup>84</sup>

The previous articulations of the male gaze continue to apply to other forms of advertising and media—particularly on the Internet—and the pornification of women on the web is an expression of racist and sexist hierarchies. When these images are present, White women are the

norm, and Black women are overrepresented, while Latinas are underrepresented.<sup>85</sup> Tracey A. Gardner characterizes the problematic characterizations of African American women in pornographic media by suggesting that "pornography capitalizes on the underlying historical myths surrounding and oppressing people of color in this country which makes it racist."<sup>86</sup> These characterizations translate from old media representations to new media forms. Structural inequalities of society are being reproduced on the Internet, and the quest for a race-, gender-, and class-less cyberspace could only "perpetuate and reinforce current systems of domination."<sup>87</sup>

More than fifteen years later, the present research corroborates these concerns. Women, particularly of color, are represented in search queries against the backdrop of a White male gaze that functions as the dominant paradigm on the Internet in the United States. The Black studies and critical Whiteness scholar George Lipsitz, of the University of California, Santa Barbara, highlights the "possessive investment in Whiteness" and the ways that the American construction of Whiteness is more "nonracial" or null. Whiteness is more than a legal abstraction formulated to conceptualize and codify notions of the "Negro," "Black Codes," or the racialization of diverse groups of African peoples under the brutality of slavery—it is an imagined and constructed community uniting ethnically diverse European Americans. Through cultural agreements about who subtly and explicitly constitutes "the other" in traditional media and entertainment such as minstrel shows, racist films and television shows produced in Hollywood, and Wild West narratives, Whiteness consolidated itself "through inscribed appeals to the solidarity of White supremacy."<sup>88</sup> The cultural practices of our society—which I argue include representations on the Internet—are part of the ways in which race-neutral narratives have increased investments in Whiteness. Lipsitz argues it this way:

As long as we define social life as the sum total of conscious and deliberate individual activities, then only *individual* manifestations of personal prejudice and hostility will be seen as racist. Systemic, collective, and coordinated behavior disappears from sight. Collective exercises of group power relentlessly channeling rewards, resources, and opportunities from one group to another will not appear to be "racist" from this perspective

because they rarely announce their intention to discriminate against individuals. But they work to construct racial identities by giving people of different races vastly different life chances.<sup>89</sup>

Consistent with trying to make sense of the ways that racial order is built, maintained, and made difficult to parse, Charles Mills, in his canonical work, *The Racial Contract*, put it this way:

One could say then, as a general rule, that white *misunderstanding, misrepresentation, evasion, and self-deception on matters related to race* are among the most pervasive mental phenomena of the past few hundred years, a cognitive and moral economy psychically required for conquest, colonization and enslavement. And these phenomena are in no way *accidental, but prescribed by the Racial Contract*, which requires a certain schedule of structured blindness and opacities in order to establish and maintain the white polity.<sup>90</sup>

This, then, is a challenge, because in the face of rampant denial in Silicon Valley about the impact of its technologies on racialized people, it becomes difficult to foster an understanding and appropriate intervention into its practices. Group identity as invoked by keyword searches reveals this profound power differential that is reflected in contemporary U.S. social, political, and economic life. It underscores how much engineers have control over the mechanics of sense making on the web about complex phenomena. It begs the question that if the Internet is a tool for progress and advancement, as has been argued by many media scholars, then cui bono—to whose benefit is it, and who holds the power to shape it? Tracing these historical constructions of race and gender offline provides more information about the context in which technological objects such as commercial search engines function as an expression of a series of social, political, and economic relations—which most of Silicon Valley's leadership is unwilling to engage with or take up.<sup>91</sup>

Studying Google keyword searches on identity, and their results, helps further thinking about what this means in relationship to marginalized groups in the United States. I take up the communications

scholar Norman Fairclough's rationale for doing this kind of critique of the discourses that contribute to the meaning-making process as a form of "critical social science."<sup>92</sup> To contextualize my method and its appropriateness to my theoretical approach, I note here that scholars who work in critical race theory and Black feminism often use a qualitative method such as close reading, which provides more than numbers to explain results and which focuses instead on the material conditions on which these results are predicated.

### Challenging Cybertopias

All of this leads to more discussion about ideologies that serve to stabilize and normalize the notion of commercial search, including the still-popular and ever-persistent dominant narratives about the neutrality and objectivity of the Internet itself—beyond Google and beyond utopian visions of computer software and hardware. The early cyberarian John Perry Barlow's infamous "A Declaration of the Independence of Cyberspace" argued in part, "We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth. We are creating a world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity."<sup>93</sup> Yet the web is not only an intangible space; it is also a physical space made of brick, mortar, metal trailers, electronics containing magnetic and optical media, and fiber infrastructure. It is wholly material in all of its qualities, and our experiences with it are as real as any other aspect of life. Access to it is predicated on telecommunications companies, broadband providers, and Internet service providers (ISPs). Its users live on Earth in myriad human conditions that make them anything but immune from privilege and prejudice, and human participation in the web is mediated by a host of social, political, and economic access points—both locally in the United States and globally.<sup>94</sup>

Since Barlow's declaration, many scholars have challenged the utopian ideals associated with the rise of the Internet and its ability to free us, such as those espoused by Barlow, linking them to neoliberal notions of individualism, personal freedom, and individual control. These linkages are important markers of the shift from public- or state-sponsored

institutions, including information institutions, as the arbiters of social freedoms to the idea that free markets, corporations, and individualized pursuits should serve as the locus of social organization. These ideas are historically rooted in notions of the universal human being, unmarked by difference, that serve as the framework for a specific tradition of thinking about *individual* pursuits of equality. Nancy Leys Stepan of Cornell University aptly describes an enduring feature of the past 270 years of liberal individualism, reinvoked by Enlightenment thinkers during the rising period of modern capitalism:

Starting in the seventeenth century, and culminating in the writings of the new social contract philosophers of the eighteenth century, a new concept of the political individual was formulated—an abstract and innovative concept, an apparent oxymoron—the imagined *universal individual* who was the bearer of equal political rights. The genius of this concept, which opened the door to the modern polis, was that it defined at least theoretically, an individual being who could be imagined so stripped of individual substantiation and specification (his unique self), that he could stand for every man. Unmarked by the myriad specificities (e.g., of wealth, rank, education, age, sex) that make each person unique, one could imagine an abstract, non-specific individual who expressed a common psyche and political humanity.<sup>95</sup>

Of course, these notions have been consistently challenged, yet they still serve as the basis for beliefs in an ideal of an unmarked humanity—nonracialized, nongendered, and without class distinction—as the final goal of human transcendence. This teleology of the abstracted individual is challenged by the inevitability of such markers and the ways that the individual particularities they signal afford differential realities and struggles, as well as privileges and possibilities. Those who become “marked” by race, gender, or sexuality as other are deviations from the universal human—they are often lauded for “transcending” their markers—while others attempt to “not see color” in a failing quest for colorblindness. The pretext of universal humanity is never challenged, and the default and idealized human condition is unencumbered by racial and gender distinction. This subtext is an important part of the narrative that somehow personal liberties can be realized through

technology because of its ability to supposedly strip us of our specifics and make us equal. We know, of course, that nothing could be further from the truth. Just ask the women of #Gamergate<sup>96</sup> and observe the ways that racist, sexist, and homophobic comments and trolling occur every minute of every hour of every day on the web.

As I have suggested, there are many myths about the Internet, including the notion that what rises to the top of the information pile is *strictly* what is most popular as indicated by hyperlinking. Were that even true, what is most popular is not necessarily what is *most true*. It is on this basis that I contend there is work to be done to contextualize and reveal the many ways that Black women are embedded within the most popular commercial search engine—Google Search—and that this embeddedness warrants an exploration into the complexities of whether the content surfaced is a result of popularity, credibility, commerciality, or even a combination thereof. Using the flawed logic of democracy in web rankings, the outcome of the searches I conducted would suggest that both sexism and pornography are the most “popular” values on the Internet when it comes to women, especially women and girls of color. In reality, there is more to result ranking than just how we “vote” with our clicks, and various expressions of sexism and racism are related.

## Searching for Black Girls



practices and product designs. Indeed, the notion that lack of participation by African Americans in Silicon Valley is framed as a “pipeline issue” posits the lack of hiring Black people as a matter of people unprepared to participate, despite evidence to the contrary. Google, Facebook, and other technology giants have been called to task for this failed logic. Laura Weidman Powers of CODE2040 stated in an interview by Jessica Guynn at *USA Today*, “This narrative that nothing can be done today and so we must invest in the youth of tomorrow ignores the talents and achievements of the thousands of people in tech from underrepresented backgrounds and renders them invisible.”<sup>1</sup> Blacks and Latinos are underemployed despite the increasing numbers graduating from college with degrees in computer science.

Filling the pipeline and holding “future” Black women programmers responsible for solving the problems of racist exclusion and misrepresentation in Silicon Valley or in biased product development is not the answer. Commercial search prioritizes results predicated on a variety of factors that are anything but objective or value-free. Indeed, there are infinite possibilities for other ways of designing access to knowledge and information, but the lack of attention to the kind of White and Asian male dominance that Guynn reported sidesteps those who are responsible for these companies’ current technology designers and their troublesome products. Few voices of African American women innovators and tech-company leaders in Silicon Valley have emerged to reframe the “diversity problems” that keep African American women at bay. One essay that grabbed the attention of many people, written for *Recode* by Heather Hiles, the former CEO of an educational technology e-portfolio company, Pathbrite, spoke directly to the limits for Black women in Silicon Valley:

I’m writing this post from the Austin airport, headed home to Oakland from SXSW. Before pulling out my laptop to compose this, I read a post on Medium that named me as one of three black women known to have raised millions in venture capital. The article began with the startling fact that less than .1 percent of venture capital in the United States is invested in black women founders. I’m not sure what sub-percentage of these are women in tech, but it doesn’t really matter when the overall numbers are so abysmal. The problem isn’t a lack of compelling women

On June 28, 2016, Black feminist and mainstream social media erupted with the announcement that Black Girls Code, an organization dedicated to teaching and mentoring African American girls interested in computer programming, would be moving into Google’s New York offices. The partnership was part of Google’s effort to spend \$150 million on diversity programs that could create a pipeline of talent into Silicon Valley and the tech industries. But just two years before, searching on “black girls” surfaced “Black Booty on the Beach” and “Sugary Black Pussy” to the first page of Google results, out of the trillions of web-indexed pages that Google Search crawls. In part, the intervention of teaching computer code to African American girls through projects such as Black Girls Code is designed to ensure fuller participation in the design of software and to remedy persistent exclusion. The logic of new pipeline investments in youth was touted as an opportunity to foster an empowered vision for Black women’s participation in Silicon Valley industries. Discourses of creativity, cultural context, and freedom are fundamental narratives that drive the coding gap, or the new coding divide, of the twenty-first century.

Part of the ethos of engaging African American women and girls in this initiative is about moving the narrative from African Americans as digitally divided to digitally *undivided*. In this framing, Black women are the targets of a variety of neoliberal science, technology, and digital innovation programs. Neoliberalism has emerged and served as a framework for developing social and economic policy in the interest of elites, while simultaneously crafting a new worldview: an ideology of individual freedoms that foreground personal creativity, contribution, and participation, as if these engagements are not interconnected to broader labor practices of systemic and structural exclusion. In the case of Google’s history of racist bias in search, no linkages are made between Black Girls Code and remedies to the company’s current employment

of color to invest in; it's a system in Silicon Valley that isn't set up to develop, encourage and create pathways for blacks, Latinos or women. Don't just take my word for it—listen to industry leaders interviewed for a USA Today story on the Valley's lack of commitment to diversity. Jessica Gwynn reports that "venture capitalists tell [Mitch Kapor] all the time that they are 'color blind' when funding companies. He's not sure they are ready to let go of a deeply rooted sense that Silicon Valley is a meritocracy."<sup>2</sup>

Hiles goes on to discuss the exclusionary practices of Silicon Valley, challenging the notion that merit and opportunity go to the smartest people prepared to innovate. Despite her being the only openly gay Black woman to raise \$12 million in venture capital for her company, she still faces tremendous obstacles that her non-Black counterparts do not. By rendering people of color as nontechnical, the domain of technology "belongs" to Whites and reinforces problematic conceptions of African Americans.<sup>3</sup> This is only exacerbated by framing the problems as "pipeline" issues instead of as an issue of racism and sexism, which extends from employment practices to product design. "Black girls need to learn how to code" is an excuse for not addressing the persistent marginalization of Black women in Silicon Valley.

### Who Is Responsible for the Results?

As a result of the lack of African Americans and people with deeper knowledge of the sordid history of racism and sexism working in Silicon Valley, products are designed with a lack of careful analysis about their potential impact on a diverse array of people. If Google software engineers are not responsible for the design of their algorithms, then who is? These are the details of what a search for "black girls" would yield for many years, despite that the words "porn," "pornography," or "sex" were not included in the search box. In the text for the first page of results, for example, the word "pussy," as a noun, is used four times to describe Black girls. Other words in the lines of text on the first page include "sugary" (two times), "hairy" (one), "sex" (one), "booty/ass" (two), "teen" (one), "big" (one), "porn star" (one), "hot" (one), "hard-core" (one), "action" (one), "galleries [sic]" (one).

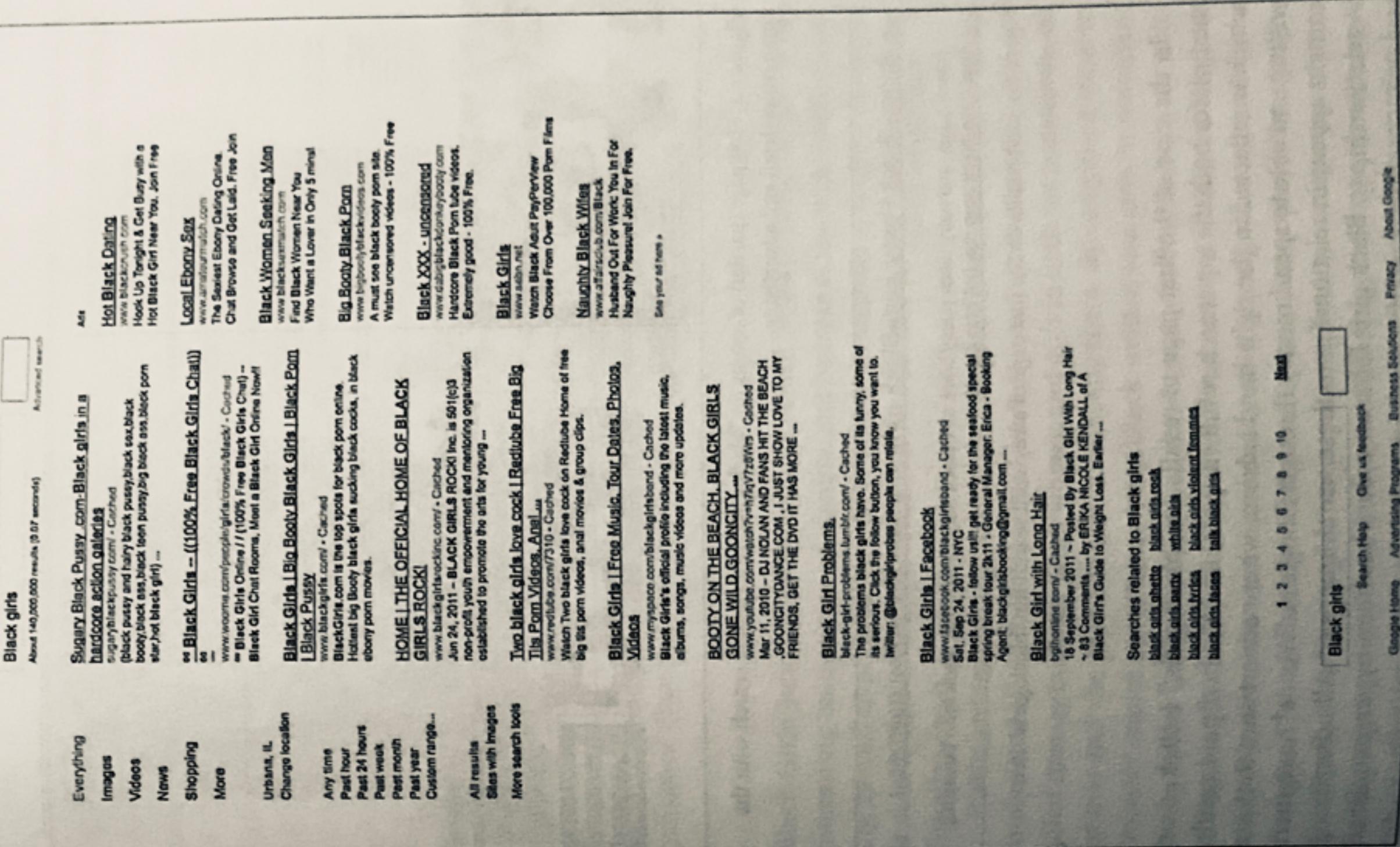


Figure 2.1. First page of search results on keywords "black girls," September 18, 2011.